IRBM 43 (2022) 325–332

Contents lists available at ScienceDirect

IRBM

www.elsevier.com/locate/irbm

Elsevier Masson France

EM|consulte

www.em-consulte.com

Original Article

ECG Data Compression Using Modified Run Length Encoding of Wavelet Coefficients for Holter Monitoring

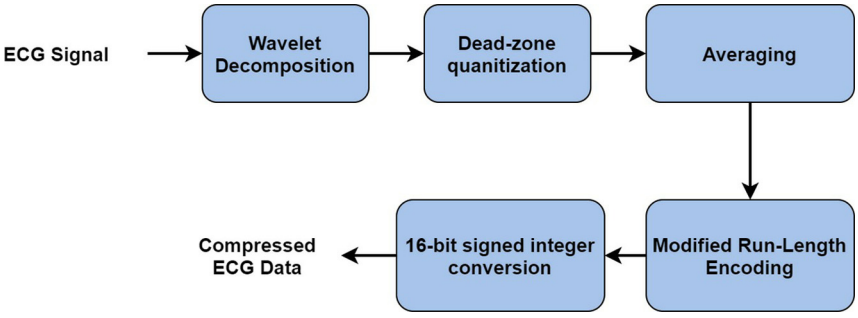
M.H. Kolekar [a], C.K. Jha [b,*], P. Kumar [a]

[a] Dept. of Electrical Engineering, Indian Institute of Technology Patna, 801106, India
[b] Dept. of Electronics and Communication Engineering, Indian Institute of Information Technology Bhagalpur, 813210, India

HIGHLIGHTS

- The proposed compression technique is based on modified run-length encoding of wavelet coefficients.
- Wavelet decomposition packs maximum energy of the signal to few transform coefficients.
- Dead-zone quantization discards small valued transform coefficients.
- Modified run-length encoding offers higher compression ratio without any information loss.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 11 January 2021
Received in revised form 31 May 2021
Accepted 3 October 2021
Available online 9 October 2021

Keywords:
ECG
Compression
Wavelet-transform
Run-length encoding
Dead-zone quantization

ABSTRACT

Objective: In cardiac patient-care, compression of long-term ECG data is essential to minimize the data storage requirement and transmission cost. Hence, this paper presents a novel electrocardiogram data compression technique which utilizes modified run-length encoding of wavelet coefficients.
Method: First, wavelet transform is applied to the ECG data which decomposes it and packs maximum energy to less number of transform coefficients. The wavelet transform coefficients are quantized using dead-zone quantization. It discards small valued coefficients lying in the dead-zone interval while other coefficients are kept at the formulated quantized output interval. Among all the quantized coefficients, an average value is assigned to those coefficients for which energy packing efficiency is less than 99.99%. The obtained coefficients are encoded using modified run-length coding. It offers higher compression ratio than conventional run-length coding without any loss of information.
Results: Compression performance of the proposed technique is evaluated using different ECG records taken from the MIT-BIH arrhythmia database. The average compression performance in terms of compression ratio, percent root mean square difference, normalized percent mean square difference, and signal to noise ratio are 17.18, 3.92, 6.36, and 28.27 dB respectively for 48 ECG records.
Conclusion: The compression results obtained by the proposed technique is better than techniques recently introduced by others. The proposed technique can be utilized for compression of ECG records of Holter monitoring.

© 2021 AGBM. Published by Elsevier Masson SAS. All rights reserved.

* Corresponding author.
E-mail addresses: mahesh@iitp.ac.in (M.H. Kolekar), ckjha.ece@iiitbh.ac.in, ckjha.bit@gmail.com (C.K. Jha), praneetk128@gmail.com (P. Kumar).

https://doi.org/10.1016/j.irbm.2021.10.001
1959-0318/© 2021 AGBM. Published by Elsevier Masson SAS. All rights reserved.

1. Introduction

The electrocardiogram (ECG) is a graphical representation of electrical impulses generated by heart muscles. It is extensively used by cardiologists to detect a multitude of cardiac diseases by examining the abnormalities in shape and pattern of the signal. Most of the time, the symptoms of a cardiac patient is transient. The regular pathological ECG test might not detect any heart abnormality as it records the ECG for a very short duration. Thus, long-term monitoring (also known as Holter monitoring) is required for cardiac patients. Holter monitoring provides long run of ECG data which is used for early detection and prevention of heart diseases. It also helps the cardiologists to determine the efficacy of medicines on patients. Since, Holter monitoring requires ECG recording for 12 hours to 48 hours continuously, the issue of efficient storage and transmission of ECG data is of great concern [1], [2]. In this regard, the adoption of compression technique solve the problem of large memory space and bandwidth requirement for storage and transmission of ECG data.

Many ECG compression techniques have been developed to reduce the memory storage of signals for efficient storage and transmission. Broadly, these techniques are of two types: lossy and lossless. In lossy compression techniques, there should be no perceptible loss of information while the lossless compression allows the exact original data to be reconstructed from the compressed data. Lossy ECG data compression techniques are classified into three types [3]: direct, parameter extraction, and transform domain methods. Direct methods use operations in the time domain to remove redundancy in ECG signals. Entropy coding, amplitude zone time epoch coding (AZTEC) improved AZTEC coding, turning-point (TP) and ASCII character encoding are examples of the direct method [4], [5]. The ECG signal contains some important features such as the P-wave section, QRS-complex, and T-wave section. The parameter extraction methods extract these features to utilize beat codebook matching and long term prediction to achieve compression. This method consists of artificial neural network (ANN), vector quantization, and peak picking [6]. The transform domain methods employ linear orthogonal transformations to decompose the ECG signal. Further, the transform coefficients are encoded appropriately to get the compressed ECG data. Discrete Fourier transform (DFT) [7], discrete cosine transform (DCT) [8], [9], [1], [10], discrete wavelet transform (DWT) [11], [12], [13], [14], [15], [16], [17], [18] are the most widely used transform techniques for ECG signal compression. Various advanced techniques based on two-dimensional transform methods such as 2D-DCT [19], 2D-DWT [20], [21], [22] and singular value decomposition (SVD) [23] have been also proposed by many researchers. Empirical mode decomposition [24], [25], [15] and accuracy driven sparse model [26] are some of the recent methods in the field of ECG data compression.

In the field of ECG data compression, transform based techniques have witnessed significant growth in the past few decades. Among transform based ECG data compression approaches, wavelet transform proved its superiority over cother transform methods in the context of compression performance [27]. Wavelet transform provides time-frequency information of the signal with the involvement of translational and dilation parameters. In wavelet decomposition of signal, energy concentrates to fewer transform coefficients. This property is used for data compression. The field of wavelet-based ECG data compression is still exploratory. The transform coefficients of wavelet decomposition of ECG signal could be arranged well to achieve better compression performance. Hence this paper presents a wavelet-based ECG data compression technique which associates with dead-zone quantization and modified run-length encoding. The proposed technique uses a modified version of run-length encoding which provides better compression ratio at the cost of no information loss.

Remaining of the paper is organized as follows: Section 2 describes modified run-length encoding. Proposed compression methodology is introduced in Section 3. Section 4 describes results and discussion of this paper. Finally, Section 5 focuses on conclusion part of this paper.

2. Run-length encoding and modifications

Run-length encoding (RLE) is a very simple form of data compression in which runs of data are stored as a single data values and their counts. Here, 'count' means the number of times a data value repeats itself continuously. Run-length encoding is particularly effective in cases where data values are in long runs and form redundancy. For example, if the original data sequence is "wwwwaabbbceeeefhhrrrttkkkkk" then run-length encoded data sequence can be represented as "w4a2b3e4f1h2r3t2k5". In the encoded data sequence, at every odd place, data values are stored followed by their count at even values i.e., 'w4' means 'w' repeated '4' times. In the given example, a significant reduction in the lengths of the two messages can be observed as the length of original message is 27 units while the encoded message requires only 20 units to store.

2.1. Drawback of run-length encoding

In some cases RLE fails to perform reasonably well. For example, if the original data sequence is "wabccdeeeffghk" then the run-length encoded data sequence will be "w1a1b1c2d1e3f2g1h1k1". Here, the length of original data sequence is 14 units while the length of the encoded data sequence is 20 units. Here, the data value which is not repeated takes 2 units in the encoded data sequence which apparently defies our objective of compression. This example highlights the drawback of compression with conventional RLE for the cases where a large number of data values are present in runs of 1 as each data value.

2.2. Modification in run-length encoding

In this paper, a modification is proposed in the conventional RLE which helps to remove the redundancy of the encoded data due to presence of single valued data instances. The modification is suggested by omitting '1' at even places of the traditional run-length encoded data. In the previous example, the original data sequence is "wabccdeeeffghk" and the run-length encoded data sequence is "w1a1b1c2d1e3f2g1h1k1". After omitting "1" at even places, the new encoded data would be "wabc2de3f2ghk" which requires 13 units length to store while traditional run-length encoded data requires 20 units length. Thus, a considerable reduction in the data size is observed. At the time of decoding, no count value of a particular data instance indicates that the data instance is present with single count.

3. Proposed compression methodology

In the proposed compression technique, first, the ECG signal is decomposed using wavelet transform which packs maximum energy of the signal to few transform coefficients. These coefficients are quantized using dead-zone quantization. After quantization, an averaging operation is performed for coefficients which have energy packing efficiency less than 99%. Further, transform coefficients are coded using modified run-length encoding. Steps of the proposed compression technique are shown in Fig. 1 using a block diagram. These steps are detailed as follows:

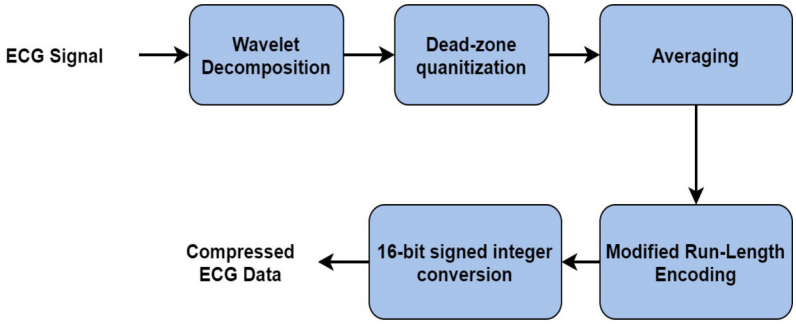


Fig. 1. Proposed compression methodology.

3.1. ECG signal

The proposed compression technique uses ECG signals of the MIT-BIH arrhythmia database which contains 48 ECG records sampled at 360 samples/ second with the resolution of 11 bits/ sample. These ECG records are acquired from 47 subjects in which 25 subjects were men aged between 32 to 89 years, and 22 subjects were women aged between 23 to 89 years.

3.2. Wavelet decomposition

The ECG signal obtained from the MIT-BIH arrhythmia database is decomposed using one dimensional wavelet transform. In the wavelet decomposition, mother wavelet of 'dabuchies' family of order 2 is used which is considered suitable for wavelet-based ECG signal compression [27]. Number of decomposition level L is chosen by the heuristically derived mathematical relation [28].

$$L = \lfloor \log_2 F_s - 2.96 \rfloor \tag{1}$$

where F_s is sampling frequency of the ECG signal. For ECG records of the MIT-BIH arrhythmia database, sampling frequency is 360 Hz. Hence, number of decomposition level used in our technique is 5.

3.3. Dead-zone quantization

Dead-zone quantization (DZQ) is employed to all obtained transform coefficients after wavelet decomposition. The DZQ scheme can be described as follows [29]:

$$D_k = \begin{cases} (-3\delta, -T_h], & if\ k = -1 \\ (-T_h, T_h), & if k = 0 \\ [T_h, 3\delta), & if k = 1 \\ [(2k-1)\delta, (2k+1)\delta), & otherwise \end{cases} \tag{2}$$

$$R_k = \begin{cases} 0, & if\ k = 0 \\ k\Delta, & for\ k = \pm 1, \pm 2, \pm 3, \ldots \end{cases} \tag{3}$$

where, D_k is the k^{th} decision interval for input transform coefficients and R_k is the corresponding quantized output value. Δ is the quantization step size while δ represents half of the quantization step size. In DZQ scheme, a small interval $(-T_h, T_h)$ is chosen around zero and all small transform coefficients falling in that interval are converted to zero. T_h is called the threshold value of the DZQ and the interval $(-T_h, T_h)$ is known as the dead-zone whose width is $2T_h$. The quantization step size Δ satisfies the condition: $(\Delta < 2T_h)$. In our technique, T_h value is determined by the relation: $T_h = 0.8 * (M - N)$, where M and N are the maximum and minimum value of the transform coefficients. In DZQ, except dead-zone and adjacent intervals for $k = 1$ and $k = -1$, all other decision intervals are of equal width 2δ. Step size Δ is determined by the relation $\Delta = \alpha T_h$. From [29], the prescribed range of α is 1.20 to 1.80. In the proposed technique, $\alpha = 1.60$ is the satisfactory choice which offers good compression performance.

3.4. Averaging of transform coefficients

After quantization, energy packing efficiency (EPE) [30] is determined for the transform coefficients. EPE can be defined as:

$$EPE = \frac{E_q}{E} \times 100\% \tag{4}$$

where E and E_q denote energy of the transform coefficients before and after quantization respectively. By the analysis of EPE of quantized transform coefficients, it is observed that first 20-50% coefficients have EPE above or equal to 99% while others have less than 99%. All the transform coefficients having EPE less than 99% are substituted by their average value. Here, the average value of coefficients having EPE less than 99% is taken. It improves the repetition of data instances which is exploited by the run-length encoding.

3.5. Modified run-length encoding

After averaging, all the transform coefficients are rounded-off to nearest integer values. In this data sequence, repetition of data instances forms redundancy. Run-length encoding (RLE) is used to remove redundancy and reduce the data size. Since conventional RLE fails considerably when the data sequence contains many data values occurring in runs of 1. Therefore, modified run-length encoding (MRLE) as described in Section 2 is proposed in this paper which helps to eliminate the drawback of RLE and offers more reduced data size. Fig. 2 illustrates steps of MRLE to transform coefficients. For analysis of theses steps, let, the data sequence of the transform coefficients is: 22223333444999997888111445792222786 53. After run-length coding, the encoded data will be: 2 4 3 4 4 3 9 5 8 3 1 3 4 2 5 1 7 1 9 1 2 4 7 1 8 1 6 1 5 1 3 1. Here, data values are placed at odd places while at even places their count values are mentioned. Now, 0.1 value is added at odd places (data instances mentioned in bracket) which results the following data: (2.1) 4 (3.1) 4 (4.1) 3 (9.1) 5 (8.1) 3 (1.1) 3 (4.1) 2 (5.1) 1 (7.1) 1 (9.1) 1 (2.1) 4 (7.1) 1 (8.1) 1 (6.1) 1 (5.1) 1 (3.1) 1. Now, count values "1" are removed which results the data: (2.1) 4 (3.1) 4 (4.1) 3 (9.1) 5 (8.1) 3 (1.1) 3 (4.1) 2 (5.1) (7.1) (9.1) (2.1) 4 (7.1) (8.1) (6.1) (5.1) (3.1). In this data sequence, if no integer value mentioned after the particular data instance, it implies that the particular data instance has the count value 1. Thus, adding the "0.1" value to the odd indexed data instances provides an easy identification of data instances for which count values are one. Instead of the "0.1" value, other values such as "0.2" or "0.3" etc. can be taken. But, the "0.1" value requires a fewer number of bits to save the data. Hence "0.1" value is se-

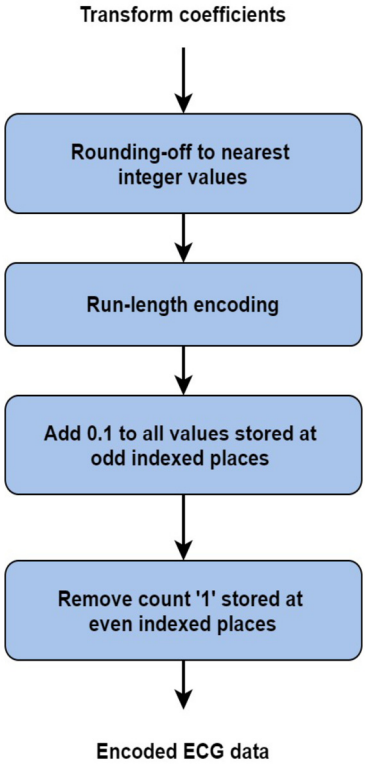

Transform coefficients
Rounding-off to nearest integer values
Run-length encoding
Add 0.1 to all values stored at odd indexed places
Remove count '1' stored at even indexed places
Encoded ECG data

Fig. 2. Steps of the Modified RLE.

lected. Thus, the data is encoded using MRLE which reduce the data size better than conventional RLE.

3.6. 16-bit signed integer conversion

The encoded ECG data instances are multiplied by 10 (because 0.1 was added at odd places) and saved as 16-bit signed integer format to reduce the number of bits required to store the data. The obtained data is the compressed form of ECG signal.

The ECG signal is reconstructed from the compressed data by following the inverse process of compression. For reconstruction of the ECG signal, the data is divided by 10 and count values are mentioned for all the data instances. After this step, 0.1 values are subtracted from the odd indexed samples. Now, inverse RLE is applied to this data which gives the original sequence of transform coefficients. Applying inverse wavelet transform to these transform coefficients results the reconstructed ECG signal. After reconstruction, some mathematical metrics are used to evaluate performance of the proposed compression technique.

3.7. Performance metrics

Performance of the proposed compression technique is evaluated using following performance metrics [1]:

3.7.1. Compression ratio (CR)

$$CR = \frac{X_{data}}{Y_{data}} \tag{5}$$

where X_{data} and Y_{data} are original and compressed ECG data size in bytes.

3.7.2. Percent root mean square difference (PRD)

$$PRD = \sqrt{\frac{\sum_{i=1}^{N}(x_i - x_i^r)^2}{\sum_{i=1}^{N} x_i^2}} \times 100 \tag{6}$$

where, x_i and x_i^r represent original and reconstructed ECG signals respectively.

3.8. Normalized PRD (PRD_1)

$$PRD_1 = \sqrt{\frac{\sum_{i=1}^{N}(x_i - x_i^r)^2}{\sum_{i=1}^{N}(x_i^2 - \overline{X})^2}} \times 100 \tag{7}$$

where $\overline{X}$ is mean value of the ECG signal.

3.8.1. Quality score (QS)

There are two forms of quality scores:.

$$QS = \frac{CR}{PRD} \tag{8}$$

$$QS = \frac{CR}{PRD_1}. \tag{9}$$

3.8.2. Signal to noise ratio (SNR)

$$SNR = 10 \times \log\left(\frac{\sum_{i=1}^{N}(x_i - \overline{X})^2}{\sum_{i=1}^{N}(x_i - x_i^r)^2}\right) \tag{10}$$

It is measured in decibel (dB) scale.

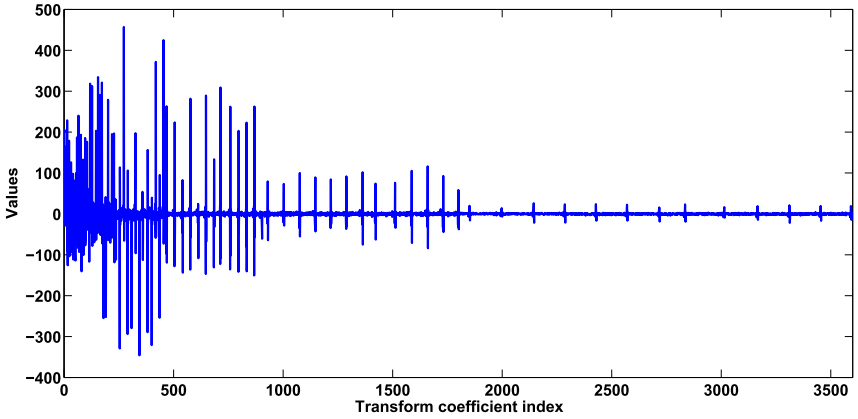
4. Results and discussion

The proposed compression technique is implemented using MATLAB 2014a software installed in a PC (intel core i3 processor). For experimentation, all 48 ECG records are taken from the MIT-BIH arrhythmia database. ECG record number 100 and 117 are chosen for visual analysis of the performance of the proposed technique. Fig. 3 illustrates transform coefficients and their values obtained after wavelet decomposition of ECG record number 100 and 117. In Fig. 3, it is clear that maximum energy of the signal squeezes to fewer transform coefficients. It verifies the energy compaction property of the wavelet transform which plays a major role in data compression.

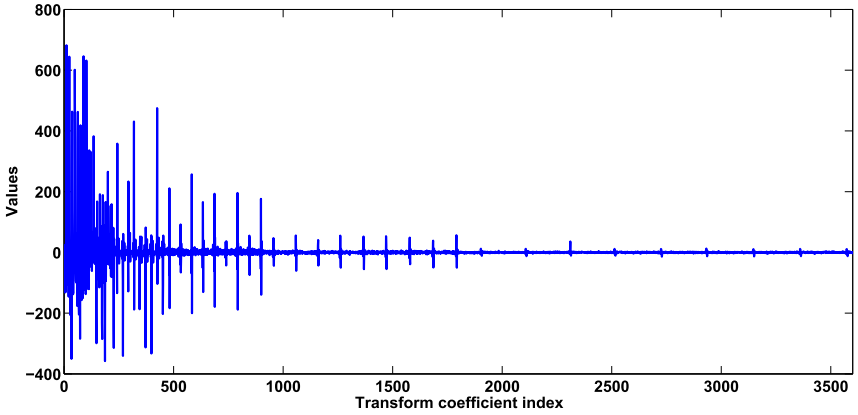
Table 1 shows the comparison of compression results with conventional RLE and MRLE adopted in the proposed technique. In the comparison, it is found that the MRLE offers improved CR without any deterioration of performance in terms of PRD, PRD^1 and SNR. In adoption of the modified RLE, there is no loss of information of the signal. Since, SNR, PRD and PRD^1 are the performance metrics which measure the distortion of the signal due to loss of information, hence values of these metrics are same with conventional and modified RLE as presented in Table 1. The proposed technique offers higher QS and QS^1 than conventional RLE which is desirable.

In Table 1, it is observed that the proposed compression technique (with MRLE) offers higher CR for ECG records no. 121 while it lowest for ECG record no. 203. PRD is higher for ECG record no. 202 while it is lowest for ECG record no. 217. Similarly, variations in other performance metrics are also observed. These performance variations are due to the morphological difference of different ECG signals. The average compression performance for 48 ECG records are 17.18, 3.92, 6.36, 4.37, 2.81 and 28.27 dB in terms CR, PRD, PRD^1, QS, QS^1 and SNR respectively. From observation of overall compression results, it is found that the proposed algorithm performs reasonably well in the compression of ECG signals.

Fig. 4 depicts the bar-chart of the average compression performance metrics with their standard deviation values. Fig. 5 shows the average CR-PRD curve for the proposed method. In this curve, it is observed that as CR increases PRD also increases. Since, the proposed compressed technique is lossy, it offers more CR at the cost of more information loss which results in more distortion of the signal. Hence, PRD increases with the increase of CR. Original



(a) Transform coefficients values after wavelet decomposition of ECG record no.100



(b) Transform coefficients values after wavelet decomposition of ECG record no.117

Fig. 3. Variation of transform coefficients values after wavelet decomposition of ECG record nos. 100 and 117.

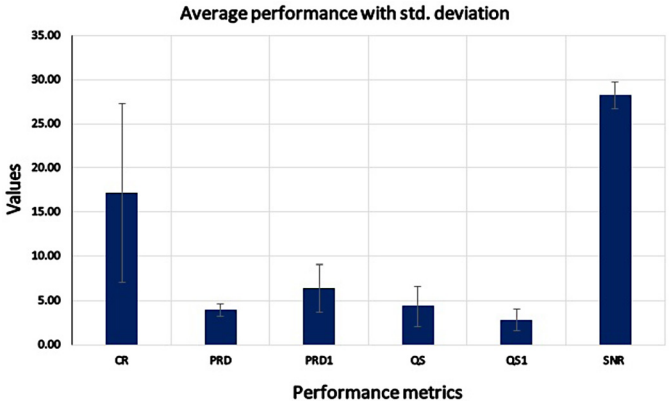


Fig. 4. Average compression performance with std. deviation.

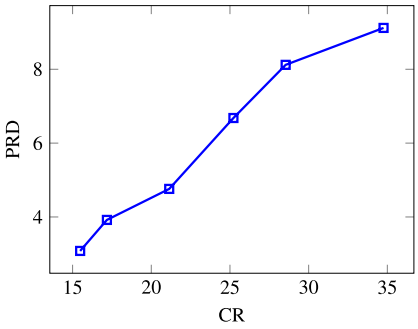


Fig. 5. Average CR-PRD curve.

and reconstructed ECG signals for ECG record numbers 100 and 117 are shown in Fig. 6. By visual inspection, it is noticed that reconstructed signals are very much similar to original ECG signals.

Table 2 illustrates the comparison of average compression performance of the proposed technique with some existing techniques based on wavelet transform, DCT, DCT based discrete orthogonal Stockwell Transform (DOST), and 2D-DCT. For the proposed compression method, CR and QS values are higher than other methods as mentioned in the comparison Table 2. In the comparison of PRD, it is observed that PRD value offered by the proposed method is a little bit higher than some other methods. In this case, QS value is taken as a criterion to choose the best compression method as per the reference [1]. Since QS value offered by the proposed method is higher than others, hence compression performance of the proposed method is better than other methods.

In addition to this, average compression performance of the proposed technique is also compared with the performance of the method in which Huffman encoding is coupled with wavelet transform followed by DZQ and averaging operations. It is shown in Table 3. It is observed that performance of the proposed technique is better in terms of CR, QS while offered PRD is same in both the cases (because modified RLE and Huffman encoding both are lossless encoding techniques).

In the proposed compression technique, wavelet decomposition, DZQ, and averaging operations are applied to the ECG data which increases the repetition of data instances. Further, this repetition is exploited using MRLE which also removes the data redundancy

Table 1
Comparison of compression results with conventional RLE and MRLE.

ECG Record No.	Compression performance with conventional RLE						Compression performance with MRLE					
	CR	PRD	PRD1	QS	QS1	SNR (dB)	CR	PRD	PRD1	QS	QS1	SNR (dB)
100	8.24	3.66	7.8	2.25	1.06	28.72	13.71	3.66	7.80	3.75	1.76	28.72
101	6.45	3.51	6.52	1.84	0.99	29.09	10.75	3.51	6.52	3.06	1.65	29.09
102	7.06	4.05	6.27	1.74	1.12	27.86	11.52	4.05	6.27	2.84	1.84	27.86
103	9.34	3.83	4.85	2.44	1.93	28.33	15.52	3.83	4.85	4.05	3.20	28.33
104	8.85	4.00	5.28	2.21	1.68	27.96	14.81	4.00	5.28	3.70	2.80	27.96
105	10.19	4.28	5.44	2.38	1.87	27.38	17.1	4.28	5.44	4.00	3.14	27.38
106	8.05	4.7	5.22	1.71	1.54	26.55	13.08	4.70	5.22	2.78	2.51	26.55
107	10.02	3.41	3.62	2.94	2.77	29.35	17.1	3.41	3.62	5.01	4.72	29.35
108	16.23	5.28	11.47	3.07	1.41	25.54	27.32	5.28	11.47	5.17	2.38	25.54
109	15.98	5.34	7.86	2.99	2.03	25.46	27.27	5.34	7.86	5.11	3.47	25.46
111	4.30	3.33	4.19	1.29	1.03	29.56	8.00	3.33	4.19	2.40	1.91	29.56
112	11.72	3.18	14.37	3.69	0.82	29.96	20.54	3.18	14.37	6.46	1.43	29.96
113	14.27	4.22	4.92	3.38	2.90	27.49	23.61	4.22	4.92	5.59	4.80	27.49
114	11.17	4.81	4.90	2.32	2.28	26.35	17.82	4.81	4.90	3.70	3.64	26.35
115	8.56	3.01	5.73	2.84	1.49	30.44	14.46	3.01	5.73	4.80	2.52	30.44
116	15.67	4.59	8.73	3.42	1.79	26.77	26.47	4.59	8.73	5.77	3.03	26.77
117	8.99	2.95	11.14	3.05	0.81	30.60	16.14	2.95	11.14	5.47	1.45	30.6
118	9.07	3.41	10.28	2.66	0.88	29.36	16.25	3.41	10.28	4.77	1.58	29.36
119	14.71	3.67	7.46	4.01	1.97	28.72	24.83	3.67	7.46	6.77	3.33	28.72
121	39.02	4.32	15.36	9.02	2.54	27.28	63.72	4.32	15.36	14.75	4.15	27.28
122	10.12	3.38	8.63	2.99	1.17	29.41	17.52	3.38	8.63	5.18	2.03	29.41
123	8.15	3.15	9.91	2.59	0.82	30.03	14.16	3.15	9.91	4.50	1.43	30.03
124	22.33	3.64	8.53	6.13	2.62	28.77	37.31	3.64	8.53	10.25	4.37	28.77
200	9.65	4.23	4.44	2.28	2.17	27.46	16.23	4.23	4.44	3.84	3.66	27.46
201	7.60	4.14	5.65	1.83	1.34	27.65	12.96	4.14	5.65	3.13	2.29	27.65
202	14.33	6.08	6.09	2.36	2.35	24.32	23.15	6.08	6.09	3.81	3.80	24.32
203	2.68	3.78	3.91	0.71	0.68	28.44	4.86	3.78	3.91	1.29	1.24	28.44
205	6.55	3.22	7.22	2.04	0.91	29.85	11.52	3.22	7.22	3.58	1.60	29.85
207	25.40	4.85	5.88	5.24	4.32	26.29	42.11	4.85	5.88	8.68	7.16	26.29
208	9.62	3.73	3.81	2.58	2.52	28.57	16.13	3.73	3.81	4.32	4.23	28.57
209	3.81	3.72	4.25	1.02	0.90	28.59	6.52	3.72	4.25	1.75	1.53	28.59
210	5.31	3.61	4.07	1.47	1.30	28.85	9.27	3.61	4.07	2.57	2.28	28.85
212	10.79	5.8	6.46	1.86	1.67	24.72	18.41	5.8	6.46	3.17	2.85	24.72
213	7.80	3.41	4.41	2.29	1.77	29.33	13.55	3.41	4.41	3.97	3.07	29.33
214	11.38	3.91	4.01	2.91	2.84	28.15	18.97	3.91	4.01	4.85	4.73	28.15
215	6.43	5.33	6.29	1.21	1.02	25.46	10.72	5.33	6.29	2.01	1.70	25.46
217	7.22	3.39	3.47	2.13	2.08	29.39	11.98	3.39	3.47	3.53	3.45	29.39
219	9.42	3.3	5.52	2.85	1.7	29.62	16.38	3.3	5.52	4.96	2.97	29.62
220	11.33	3.27	7.04	3.47	1.61	29.71	19.17	3.27	7.04	5.86	2.72	29.71
221	4.77	3.2	3.63	1.49	1.31	29.89	8.33	3.2	3.63	2.60	2.29	29.89
222	3.43	3.72	4.73	0.92	0.73	28.59	5.87	3.72	4.73	1.58	1.24	28.59
223	8.94	3.43	6.00	2.61	1.49	29.29	15.43	3.43	6.00	4.50	2.57	29.29
228	4.58	3.25	5.71	1.41	0.8	29.76	8.6	3.25	5.71	2.65	1.51	29.76
230	11.55	4.11	5.26	2.81	2.2	27.72	19.57	4.11	5.26	4.76	3.72	27.72
231	7.19	3.97	4.17	1.81	1.72	28.02	12.19	3.97	4.17	3.07	2.92	28.02
232	4.71	3.83	7.39	1.23	0.64	28.33	8.46	3.83	7.39	2.21	1.14	28.33
233	8.00	3.3	3.54	2.42	2.26	29.64	13.94	3.30	3.54	4.22	3.94	29.64
234	6.71	3.76	3.91	1.78	1.72	28.49	11.52	3.76	3.91	3.06	2.95	28.49
Average	10.16	3.92	6.36	2.58	1.66	28.27	17.18	3.92	6.36	4.37	2.81	28.27

Table 2
Comparison of compression performance in terms of CR, PRD and QS.

Techniques	CR	PRD	QS
Beta Wavelet [17]	5.67	2.40	2.36
DCT [10]	11.49	3.43	3.82
DCT based DOST [31]	6.27	5.37	1.49
2D-DCT [19]	12.00	6.35	1.89
DWT [11]	14.2	3.84	3.69
Proposed	**17.18**	**3.92**	**4.37**

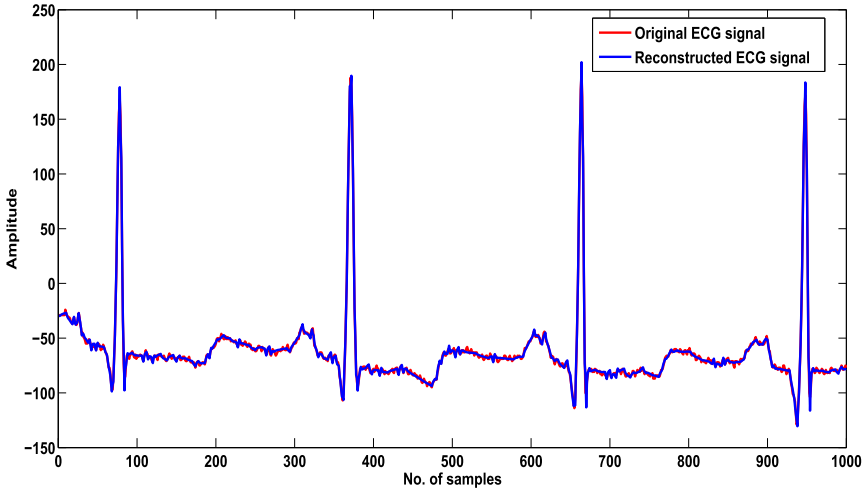
Table 3
Performance comparison with Huffman encoding.

Technique	CR	PRD	QS
Huffman	10.73	3.92	2.73
Modified RLE	17.18	3.92	4.37

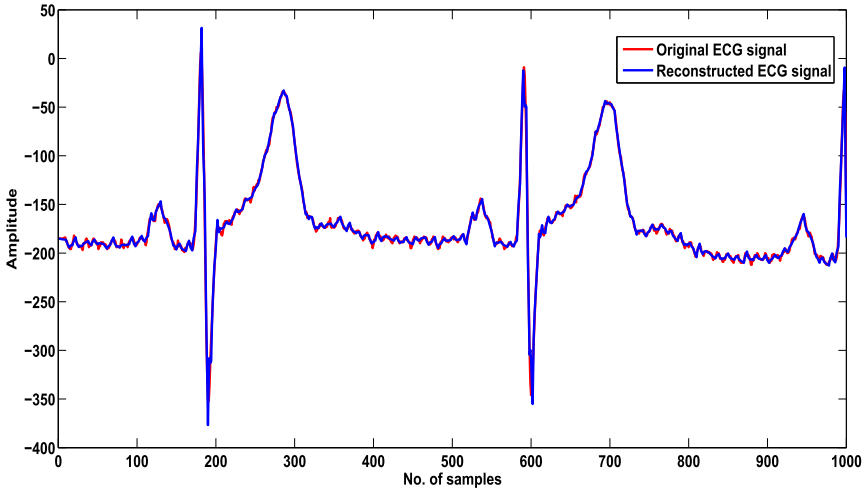
due to the single-valued data instances produced by conventional run-length encoding. Hence, MRLE produces good compression results for ECG data when it is combined with the earlier mentioned processing steps which are lossy. In future, these steps can be replaced by some other technique to reduce the loss of information or to get the lossless compression.

5. Conclusion

This paper presents an ECG data compression technique based on modified run-length encoding of wavelet transform coefficients. Original ECG signal is decomposed using wavelet transform up to 5^{th} decomposition level with 'db2' wavelet family. Further, dead-zone quantization, averaging, modified run-length encoding and 16-bit signed integer conversion of coefficients are associated to the proposed technique to achieve good compression. The proposed compression technique is evaluated using 48 ECG records of the MIT-BIH arrhythmia database. Average CR, PRD, PRD^1, QS, QS^1 and SNR obtained for theses ECG records are 17.18, 3.92, 6.36, 4.37, 2.81 and 28.27 respectively. Comparison of compression results



(a) ECG record no.100



(b) ECG record no.117

Fig. 6. Original and reconstructed signals.

with the existing techniques shows that the proposed technique offer better performance than others.

Human and animal rights

The authors declare that the work described has not involved experimentation on humans or animals.

Disclosure of interest

The authors declare that they have no known competing financial or personal relationships that could be viewed as influencing the work reported in this paper.

Author contributions

All authors attest that they meet the current International Committee of Medical Journal Editors (ICMJE) criteria for Authorship.

CRediT authorship contribution statement

M.H. Kolekar: Conceptualization, Formal analysis, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – review & editing. **C.K. Jha:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Writing – original draft. **P. Kumar:** Conceptualization, Formal analysis, Methodology, Software, Validation, Writing – original draft.

Declaration of competing interest

The authors declare that they have no conflicts of interest.

References

[1] Lee S, Kim J, Lee M. A real-time ECG data compression and transmission algorithm for an e-health device. IEEE Trans Biomed Eng 2011;58(9):2448–55.
[2] Jha CK, Kolekar MH. Classification and compression of ECG signal for holter device. In: Biomedical signal and image processing in patient care. IGI Global; 2018. p. 46–63.
[3] Jalaleddine SM, Hutchens CG, Strattan RD, Coberly WA. ECG data compression techniques-a unified approach. IEEE Trans Biomed Eng 1990;37(4):329–43.
[4] Cox J, Nolle F, Fozzard H, Oliver G. AZTEC, a preprocessing program for real-time ECG rhythm analysis. IEEE Trans Biomed Eng 1968;2:128–9.
[5] Mukhopadhyay S, Mitra S, Mitra M. An ECG signal compression technique using ASCII character encoding. Measurement 2012;45(6):1651–60.
[6] Cohen A, Poluta M, Scott-Millar R. Compression of ECG signals using vector quantization. In: South African sympos on comm and signal processing; 1990. p. 49–54.
[7] Ma J, Zhang T, Dong M. A novel ECG data compression method using adaptive Fourier decomposition with security guarantee in e-health applications. IEEE J Biomed Health Inform 2015;19(3):986–94.

[8] Bendifallah A, Benzid R, Boulemden M. Improved ECG compression method using discrete cosine transform. Electron Lett 2011;47(2):87–9.
[9] Chandra S, Sharma A, Singh G. Computationally efficient cosine modulated filter bank design for ECG signal compression. IRBM 2020;41(1):2–17.
[10] Jha CK, Kolekar MH. ECG data compression algorithm for tele-monitoring of cardiac patients. Int J Telemed Clin Pract 2017;2(1):31–41.
[11] Abo-Zahhad MM, Abdel-Hamid TK, Mohamed AM. Compression of ECG signals based on DWT and exploiting the correlation between ECG signal samples. Int J Commun Netw Syst Sci 2014;7(01):53.
[12] Jha CK, Kolekar MH. Tunable Q-wavelet based ECG data compression with validation using cardiac arrhythmia patterns. Biomed Signal Process Control 2021;66:102464.
[13] Jha CK, Kolekar MH. Efficient ECG data compression and transmission algorithm for telemedicine; 2016. p. 1–6.
[14] Abo-Zahhad M, Al-Ajlouni AF, Ahmed SM, Schilling RJ. A new algorithm for the compression of ECG signals based on mother wavelet parameterization and best-threshold levels selection. Digit Signal Process 2013;23(3):1002–11.
[15] Jha CK, Kolekar MH. Empirical mode decomposition and wavelet transform based ECG data compression scheme. IRBM 2021;42(1):65–72.
[16] Jha CK, Kolekar MH. Diagnostic quality assured ECG signal compression with selection of appropriate mother wavelet for minimal distortion. IET Sci Meas Technol 2019;13(4):500–8.
[17] Kumar R, Kumar A, Pandey RK. Beta wavelet based ECG signal compression using lossless encoding with modified thresholding. Comput Electr Eng 2013;39(1):130–40.
[18] Jha CK, Kolekar MH. Classification and compression of ECG signal for holter device. Biomed Signal Image Process Patient Care 2017:46.
[19] Lee H, Buckley KM. ECG data compression using cut and align beats approach and 2-D transforms. IEEE Trans Biomed Eng 1999;46(5):556–64.
[20] Lu Z, Kim DY, Pearlman WA. Wavelet compression of ECG signals by the set partitioning in hierarchical trees algorithm. IEEE Trans Biomed Eng 2000;47(7):849–56.
[21] Tai S-C, Sun C, Yan W-C. A 2-D ECG compression method based on wavelet transform and modified SPIHT. IEEE Trans Biomed Eng 2005;52(6):999–1008.
[22] Chou H-H, Chen Y-J, Shiau Y-C, Kuo T-S. An effective and efficient compression algorithm for ECG signals with irregular periods. IEEE Trans Biomed Eng 2006;53(6):1198–205.
[23] Wei J-J, Chang C-J, Chou N-K, Jan G-J. ECG data compression using truncated singular value decomposition. IEEE Trans Inf Technol Biomed 2001;5(4):290–9.
[24] Zhao C, Chen Z, Meng J, Xiang X. Electrocardiograph compression based on sifting process of empirical mode decomposition. Electron Lett 2016;52(9):688–90.
[25] Wang X, Chen Z, Luo J, Meng J, Xu Y. ECG compression based on combining of EMD and wavelet transform. Electron Lett 2016;52(19):1588–90.
[26] Grossi G, Lanzarotti R, Lin J. High-rate compression of ECG signals by an accuracy-driven sparsity model relying on natural basis. Digit Signal Process 2015;45:96–106.
[27] Manikandan MS, Dandapat S. Wavelet-based electrocardiogram signal compression methods and their performances: a prospective review. Biomed Signal Process Control 2014;14:73–107.
[28] Al-Fahoum AS. Quality assessment of ECG compression techniques using a wavelet-based diagnostic measure. IEEE Trans Inf Technol Biomed 2006;10(1):182–91.
[29] Chen J, Wang F, Zhang Y, Shi X. ECG compression using uniform scalar dead-zone quantization and conditional entropy coding. Med Eng Phys 2008;30(4):523–30.
[30] Rajoub BA. An efficient coding algorithm for the compression of ECG signals using the wavelet transform. IEEE Trans Biomed Eng 2002;49(4):355–62.
[31] Jha CK, Kolekar MH. Electrocardiogram data compression using DCT based discrete orthogonal stockwell transform. Biomed Signal Process Control 2018;46:174–81.